1.0

1.1

1.25 1.4

45
50
55

28
3.2
36
40

2.5

2.2

2.0

1.8

1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

LEVEL II

ON GENERALIZATIONS OF COCHRAN'S THEOREM

AND PROJECTION MATRICES

BY

AKIMICHI TAKEMURA

TECHNICAL REPORT NO. 44

AUG   1980

DEPARTMENT OF STATISTICS

STANFORD   UNIVERSITY

STANFORD, CALIFORNIA

DTIC
ELECTE
DEC 5 1980
D

AD A092595

DDC FILE COPY

80 12 03 006

ON GENERALIZATIONS OF COCHRAN'S THEOREM

AND PROJECTION MATRICES

by

AKIMICHI TAKEMURA

Stanford University

TECHNICAL REPORT NO. 44

AUGUST 1980

DTIC

S **ELECTE** **D**

DEC 5 1980

D

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ON GENERALIZATIONS OF COCHRAN'S THEOREM
AND PROJECTION MATRICES

by

Akimichi Takemura

1. Introduction

In statistical analysis the decomposition of a sum of squares into
various independent components is essential in ANOVA-type procedures.
In this connection Cochran's theorem gives nice conditions to guarantee
the independence of components. There have been many generalizations
or refinements of Cochran's theorem, and recently Anderson and Styan
(1980) reviewed those results and gave new ones.

In this paper we present a unified approach to those problems by
consistent use of "projection" matrices which are generally not ortho-
gonal. This approach enables us to extend various results to much greater
generality.

In section 2 we review the properties of projection. In section 3
we prove some generalizations of Cochran's theorem. In section 4 we
discuss decomposition of matrices into projections to their eigenspaces
and its relation to Cochran-type decomposition.

Matrices may be considered as real or complex. Proofs are given
in general vector space terms, and often they cover both cases. In section
4 we use complex matrices explicitly.

1

2. Projection Matrix

In this section properties of projection matrices are summarized. To be self-contained brief proofs will be given.

A matrix $\underset{\sim}{A}$ is called a __projection__ matrix if it is idempotent, i.e., $\underset{\sim}{A}^2 = \underset{\sim}{A}$. If in addition $\underset{\sim}{A}' = \underset{\sim}{A}$ (or $\underset{\sim}{A}^* = \underset{\sim}{A}$ for complex matrices) then $\underset{\sim}{A}$ is called an orthogonal projection matrix.[1] Orthogonal projections have many extra properties which are not shared by general projections. However, there are quite a few facts which have counterparts in general projections or hold without orthogonality. Moreover, the advantage of developing a theory for general projections is that the theory holds for vector spaces __not__ equipped with an inner product.

We need some definitions about subspaces of a vector space. Let X be a vector space and $U_1, \ldots, U_k$ be subspaces of X. $U_1, \ldots, U_k$ are called __(linearly) independent__ if $\underset{\sim}{x}_i \in U_i$, $i = 1, \ldots, k$, $\sum_{i=1}^{k} \underset{\sim}{x}_i = \underset{\sim}{0}$ implies $\underset{\sim}{x}_i = \underset{\sim}{0}$ $i = 1, \ldots, k$. When $U_1, \ldots, U_k$ are independent, the subspace U spanned by $U_1, \ldots, U_k$ is called the __direct sum__ of U_i's and denoted by $U = U_1 + \ldots + U_k = \sum_{i=1}^{k} U_i$. An element $\underset{\sim}{y}$ of U has a unique representation $\underset{\sim}{y} = \underset{\sim}{x}_1 + \ldots + \underset{\sim}{x}_k$ where $\underset{\sim}{x}_i \in U_i$, $i = 1, \ldots, k$. It is easy to see that the union of bases of U_i's form a basis of U and therefore $\dim U = \sum \dim U_i$. For a matrix $\underset{\sim}{A}$ the __column vector space__ or the __range__ of $\underset{\sim}{A}$ is denoted by $C(\underset{\sim}{A})$. The __null space__ of $\underset{\sim}{A}$ is denoted by $N(\underset{\sim}{A})$. $\lambda(\underset{\sim}{A}) = \{\lambda_1, \ldots, \lambda_k\}$ is the set of distinct eigenvalues of $\underset{\sim}{A}$ and $E(\lambda_i) = N(\underset{\sim}{A} - \lambda_i \underset{\sim}{I})$ denotes the __eigenspace__ associated with λ_i.

---

[1] Usually "projection" is used to mean orthogonal projection but here we use it for general projection.

Now we state properties of projection matrices.

Proposition 2.1. The following conditions are equivalent:

(i) $\underset{\sim}{A}$ is a projection.

(ii) $\underset{\sim}{I} - \underset{\sim}{A}$ is a projection.

(iii) $C(\underset{\sim}{A}) = \{\underset{\sim}{x}: \underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{x}\}$.

(iv) $N(\underset{\sim}{I}-\underset{\sim}{A}) = C(\underset{\sim}{A})$ or $N(\underset{\sim}{A}) = C(\underset{\sim}{I}-\underset{\sim}{A})$.

(v) $C(\underset{\sim}{A})$ and $C(\underset{\sim}{I}-\underset{\sim}{A})$ are independent.

Proof. We show that (i) is equivalent to each of (ii)-(v) .

(ii): (i) $\Leftrightarrow \underset{\sim}{A}(\underset{\sim}{I}-\underset{\sim}{A}) = \underset{\sim}{0} \Leftrightarrow$ (ii).

(iii): Let $V = \{\underset{\sim}{x}: \underset{\sim}{A}\underset{\sim}{x}=\underset{\sim}{x}\}$. Assume $\underset{\sim}{A}^2 = \underset{\sim}{A}$. If $\underset{\sim}{x} \in C(\underset{\sim}{A})$ then $\underset{\sim}{x} = \underset{\sim}{A}\underset{\sim}{y} = \underset{\sim}{A}^2\underset{\sim}{y} = \underset{\sim}{A}(\underset{\sim}{A}\underset{\sim}{y}) = \underset{\sim}{A}\underset{\sim}{x}$. Therefore $\underset{\sim}{x} \in V$. If $\underset{\sim}{x} \in V$ then $\underset{\sim}{x} = \underset{\sim}{A}\underset{\sim}{x} \in C(\underset{\sim}{A})$. Therefore $V = C(\underset{\sim}{A})$.

Assume $C(\underset{\sim}{A}) = V$. Then $\underset{\sim}{A}\underset{\sim}{x} \in C(\underset{\sim}{A}) = V$. Hence $\underset{\sim}{A}(\underset{\sim}{A}\underset{\sim}{x}) = \underset{\sim}{A}\underset{\sim}{x} \ \forall\underset{\sim}{x}$. Therefore $\underset{\sim}{A}^2 = \underset{\sim}{A}$.

(iv): (iv) $\Leftrightarrow$ (iii).

(v): Necessity is proved in the next proposition. Sufficiency is proved noting

$$(\underset{\sim}{A}-\underset{\sim}{A}^2)\underset{\sim}{x} - (\underset{\sim}{I}-\underset{\sim}{A})\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{0} , \quad \forall\underset{\sim}{x} ,$$

and $(\underset{\sim}{A}-\underset{\sim}{A}^2)\underset{\sim}{x} \in C(\underset{\sim}{A})$, $(\underset{\sim}{I}-\underset{\sim}{A})\underset{\sim}{A}\underset{\sim}{x} \in C(\underset{\sim}{I}-\underset{\sim}{A})$. Q.E.D.

Properties (iii), (iv), (v) show that $\underset{\sim}{A}$ projects a vector $\underset{\sim}{x} \notin C(\underset{\sim}{A})$ to $C(\underset{\sim}{A})$ along a line parallel to $C(\underset{\sim}{I}-\underset{\sim}{A})$.

__Proposition 2.2.__ Suppose $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_k$ are projections such that

$\underset{\sim}{A}_i \underset{\sim}{A}_j = \underset{\sim}{0}$ for $i \neq j$. Then (i) $\underset{\sim}{A} = \underset{\sim}{A}_1 + \ldots + \underset{\sim}{A}_k$ is a projection, and

(ii) $C(\underset{\sim}{A}_i)$'s are independent and $C(\underset{\sim}{A}) = \sum_{i=1}^{k} C(\underset{\sim}{A}_i)$.

__Proof.__ Let $\underset{\sim}{x}_i \in C(\underset{\sim}{A}_i)$ then $\underset{\sim}{x}_i = \underset{\sim}{A}_i \underset{\sim}{y}_i$. Suppose $\sum_{i=1}^{k} \underset{\sim}{x}_i = \underset{\sim}{0}$.

Then

$$\underset{\sim}{0} = \underset{\sim}{A}_j (\sum_{i=1}^{k} \underset{\sim}{x}_i) = \underset{\sim}{A}_j^2 \underset{\sim}{y}_j = \underset{\sim}{A}_j \underset{\sim}{y}_j = \underset{\sim}{x}_j , \quad j = 1, \ldots, k .$$

Therefore $C(\underset{\sim}{A}_i)$'s are independent. $\underset{\sim}{A}\underset{\sim}{x} = \sum_{i=1}^{k} \underset{\sim}{A}_i \underset{\sim}{x}$, therefore

(2.1) $$C(\underset{\sim}{A}) \subset \sum_{i=1}^{k} C(\underset{\sim}{A}_i) .$$

On the other hand $\underset{\sim}{A}\underset{\sim}{A}_i \underset{\sim}{x} = \underset{\sim}{A}_i^2 \underset{\sim}{x} = \underset{\sim}{A}_i \underset{\sim}{x}$ implies $C(\underset{\sim}{A}_i) \subset C(\underset{\sim}{A})$, $i = 1, \ldots, k$.

Therefore

(2.2) $$C(\underset{\sim}{A}) \supset \sum_{i=1}^{k} C(\underset{\sim}{A}_i) .$$

From (2.1) and (2.2) $C(\underset{\sim}{A}) = \sum_{i=1}^{k} C(\underset{\sim}{A}_i)$. Q.E.D.

The converse of the above proposition is given by the following

proposition.

__Proposition 2.3.__ Let $\underset{\sim}{A}$ be a projection onto $V = C(\underset{\sim}{A})$. Suppose

$V = \sum_{i=1}^{k} V_i$ where V_i's are given independent subspaces. Then

$\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{A}_i$ where $\underset{\sim}{A}_i$'s are uniquely determined projections such that

$\underset{\sim}{A}_i \underset{\sim}{A}_j = \underset{\sim}{0}$ for $i \neq j$ and $C(\underset{\sim}{A}_i) = V_i$, $i = 1, \ldots, k$.

__Proof.__ Let $V_0 = C(\underset{\sim}{I} - \underset{\sim}{A})$. Then $V_0, V_1, \ldots, V_k$ are independent and

$\sum_{i=0}^{k} V_i = X$ (whole space). Now we choose vectors $\{\underset{\sim}{x}_{01}, \ldots, \underset{\sim}{x}_{0m_0}, \underset{\sim}{x}_{11}, \ldots,$

$\underset{\sim}{x}_{1m_1}, \ldots, \underset{\sim}{x}_{km_k}\}$ such that $\{\underset{\sim}{x}_{i1}, \ldots, \underset{\sim}{x}_{im_i}\}$ form a basis for V_i .

4

Then x_{ij}'s form a basis for the whole space. Let

$$T = (x_{01}, \ldots, x_{km_k}) = (T_0, T_1, \ldots, T_k)$$

be a matrix with column vectors x_{ij} with $T_i = (x_{i1}, \ldots, x_{im_i})$, $i = 0$, $1, \ldots, k$. Let T^{-1} be partitioned accordingly

$$T^{-1} = \begin{pmatrix} T^0 \\ \vdots \\ T^k \end{pmatrix}.$$

Then

$$I = TT^{-1} = \sum_{i=0}^{k} T_i T^i.$$

Let $A_i = T_i T^i$. Now it is easy to check that $A_i^2 = A_i$, $C(A_i) = V_i$, and $A_i A_j = 0$ for $i \neq j$. To prove uniqueness let $A = \sum_{i=1}^{k} A_i = \sum_{i=1}^{k} B_i$. Then $\sum_{i=1}^{k} (A_i - B_i) x = 0$, $\forall x$, implies $A_i x = B_i x$, $i = 1, \ldots, k$, because $(A_i - B_i) x \in V_i$. Hence $A_i = B_i$, $i = 1, \ldots, k$. Q.E.D.

In view of the above two propositions it may well be justified to call projection matrices A_i, $i = 1, \ldots, k$, (linearly) independent if $A_i A_j = 0$ for $i \neq j$.

We note here the relation between diagonalizable matrices and projections. Let A be a $p \times p$ diagonalizable matrix and let $\lambda(A) = \{\lambda_1, \ldots, \lambda_k\}$. There exists a nonsingular T such that

$$T^{-1} A T = D_\lambda$$

where D_λ is a diagonal matrix with λ_i's as diagonal elements in an appropriate order, i.e.,

$$D_\lambda = \text{diag}(\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_k).$$

Now $\underset{\sim}{A} = \underset{\sim\sim}{T}D_\lambda\underset{\sim}{T}^{-1}$ and with the similar partitioning of $\underset{\sim}{T}\,\underset{\sim}{T}^{-1}$ as in the proof of Proposition 2.3

$$(2.3) \qquad\qquad \underset{\sim}{A} = \sum_{i=1}^{k} \lambda_i \underset{\sim}{H}_i$$

where $\underset{\sim}{H}_i$'s are independent projections onto eigenspace $E(\lambda_i)$'s of λ_i's . It follows that for any polynomial P

$$(2.4) \qquad\qquad P(\underset{\sim}{A}) = \sum_{i=1}^{k} P(\lambda_i) \underset{\sim}{H}_i \ .$$

$\underset{\sim}{H}_i$'s are unique by Proposition 2.3 and $\underset{\sim}{I} = \sum_{i=1}^{k} \underset{\sim}{H}_i$. Conversely if $\sum_{i=1}^{k} E(\lambda_i) = X$ [2] then $\underset{\sim}{A}$ can be diagonalized by $\underset{\sim}{T}$ as in the proof of Proposition 2.3.

When $0 \in \lambda(\underset{\sim}{A})$ then putting $\lambda_k = 0$ we can write (2.3) as

$$\underset{\sim}{A} = \sum_{i=1}^{k-1} \lambda_i \underset{\sim}{H}_i \ .$$

(2.4) holds with k replaced by $k-1$ if $P(0) = 0$, i.e., P does not have constant term.

Despite its apparent plausibility we need the diagonalizability of a projection matrix to prove:

Proposition 2.4. If $\underset{\sim}{A}^2 = \underset{\sim}{A}$ then $\text{rank } \underset{\sim}{A} = \text{tr } \underset{\sim}{A}$.

Proof. $\underset{\sim}{A}^2 = \underset{\sim}{A}$ implies $\lambda(\underset{\sim}{A}) \subset \{0,1\}$. It will be shown that $\underset{\sim}{A}$ is diagonalizable in section 4 and therefore

$$\text{rank } \underset{\sim}{A} = \text{number of eignevalues } 1$$

$$= \text{tr } \underset{\sim}{A} \ . \qquad\qquad \text{Q.E.D.}$$

---

[2] Eigenspaces $E(\lambda_i)$'s are always independent.

3. Cochran-Type Theorems

In terms of matrix theory Cochran's theorem essentially asserts the following: Let $\underset{\sim}{A}$ be an orthogonal projection and $\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{A}_i$ where $\underset{\sim}{A}_i$'s are symmetric. If $\text{rank } \underset{\sim}{A} = \sum_{i=1}^{k} \text{rank } \underset{\sim}{A}_i$ then $\underset{\sim}{A}_i$'s are independent orthogonal projections. Various kinds of extensions and refinements are discussed in Anderson and Styan (1980). In this section we present simpler proofs for a number of them and give a generalization (Theorem 3.3 together with Corollary 4.2).

We begin by showing that Cochran's theorem holds <u>without orthogonality</u>.

<u>Theorem 3.1</u>. (Cochran's theorem) Let $\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{A}_i$. $\underset{\sim}{A}_i$'s are independent projections if and only if

$\qquad$ (i) $\underset{\sim}{A}$ is a projection,

$\qquad$ (ii) $\text{rank } \underset{\sim}{A} = \sum_{i=1}^{k} \text{rank } \underset{\sim}{A}_i$.

The key point to prove this is contained in

<u>Lemma 3.1</u>. Let $\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{A}_i$ where $\underset{\sim}{A}_i$'s are general matrices. Then

$\qquad$ (i) $C(\underset{\sim}{A}_i)$'s are independent and $C(\underset{\sim}{A}) = \sum_{i=1}^{k} C(\underset{\sim}{A}_i)$

if and only if

$\qquad$ (ii) $\text{rank } \underset{\sim}{A} = \sum_{i=1}^{k} \text{rank } \underset{\sim}{A}_i$.

<u>Proof</u>.

Necessity: $\dim\{C(\underset{\sim}{A})\} = \sum_{i=1}^{k} \dim\{C(\underset{\sim}{A}_i)\}$. This is equivalent to $\text{rank } \underset{\sim}{A} = \sum_{i=1}^{k} \text{rank } \underset{\sim}{A}_i$.

Sufficiency: We show the independence of $C(A_i)$ by contradiction. Suppose that there are vectors $x_1,\ldots,x_k$ with at least one nonzero vector such that $x_i \in C(A_i)$, $i = 1,\ldots,k$, and $\sum_{i=1}^{k} x_i = 0$. Without loss of generality let

$$0 \neq x_1 = -\sum_{i=2}^{k} x_i \ .$$

Let $\{x_1, q_1, \ldots, q_m\}$ be a basis for $C(A_1)$ where $m = \text{rank } A_1 - 1$. Then $C(A) \subset C(A_1,\ldots,A_k) = C(q_1,\ldots,q_m,A_2,\ldots,A_k)$. But $\dim\{C(A)\} = \text{rank } A$ and

$$\dim C(q_1,\ldots,q_m, A_2,\ldots,A_k) \leq m + \sum_{i=2}^{k} \dim C(A_i) = \text{rank } A - 1$$

which is a contradiction. Therefore $C(A_i)$'s are independent. Now $C(A) \subset \sum_{i=1}^{k} C(A_i)$ but again by the rank condition (ii) we get $C(A) = \sum_{i=1}^{k} C(A_i)$. Q.E.D.

Proof of Theorem 3.1.

Sufficiency: From the lemma we know $C(A_i) \subset C(A)$, $i = 1,\ldots,k$. Therefore by (iii) of Proposition 2.1 we get $AA_i x = A_i x$. Then

$$\sum_{j \neq i} A_j A_i x + (A_i^2 - A_i) x = 0 \ .$$

By independence of $C(A_i)$'s we get $A_j A_i x = 0$ $i \neq j$ and $(A_i^2 - A_i) x = 0$. This being true for any x shows $A_j A_i = 0$, $i \neq j$, and $A_i^2 = A_i$.

Necessity: In view of lemma 3.1 necessity was proved already in Proposition 2.2. Q.E.D.

Chipman and Rao (1964), Khatri (1968) give more algebraic proofs of this theorem.

8

We proceed to generalize this result to an r-potent matrix. A matrix $\underset{\sim}{A}$ is called <u>r-potent</u> if $\underset{\sim}{A}^r = \underset{\sim}{A}$. A general theorem is given by Anderson and Styan (1980, their theorem 3.3). Here we present a simpler (but equivalent) version.

<u>Theorem 3.2</u>. (r-potent matrix) Let $\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{A}_i$ and $r \geq 3$.

(i) $\underset{\sim}{A}_i^r = \underset{\sim}{A}_i$, $\underset{\sim}{A}_i \underset{\sim}{A}_j = \underset{\sim}{0}$, $i \neq j$,

if and only if

(i') $\underset{\sim}{A}^r = \underset{\sim}{A}$, (ii') $\text{rank } \underset{\sim}{A} = \sum_{i=1}^{k} \text{rank } \underset{\sim}{A}_i$,

(iii') $\underset{\sim}{A}\underset{\sim}{A}_i = \underset{\sim}{A}_i\underset{\sim}{A}$, $i = 1,\ldots,k$.

<u>Proof</u>.

Necessity: We need to prove (ii'). By multiplying (i) by $\underset{\sim}{A}_i^{r-2}$ or $\underset{\sim}{A}_j^{r-2}$ we get $\underset{\sim}{A}_i^{2(r-1)} = \underset{\sim}{A}_i^{r-1}$, $\underset{\sim}{A}_i^{r-1} \underset{\sim}{A}_j^{r-1} = \underset{\sim}{0}$, $i \neq j$. Therefore the previous theorem applies to $\underset{\sim}{A}_i^{r-1}$ and we get

$$\text{rank } \underset{\sim}{A}^{r-1} = \sum_{i=1}^{k} \text{rank } \underset{\sim}{A}_i^{r-1}.$$

But (i) and (i') show that $\text{rank } \underset{\sim}{A}^{r-1} = \text{rank } \underset{\sim}{A}$, $\text{rank } \underset{\sim}{A}_i^{r-1} = \text{rank } \underset{\sim}{A}_i$ noting $\text{rank } \underset{\sim}{A} \geq \text{rank } \underset{\sim}{A}^{r-1} \geq \text{rank } \underset{\sim}{A}^r = \text{rank } \underset{\sim}{A}$, etc. (ii') follows from this.

Sufficiency: From (iii') we get

$$\sum_{j=1}^{k} \underset{\sim}{A}_j \underset{\sim}{A}_i \underset{\sim}{x} = \underset{\sim}{A}_i \underset{\sim}{A} \underset{\sim}{x}.$$

Hence

$$\sum_{j \neq i} \underset{\sim}{A}_j \underset{\sim}{A}_i \underset{\sim}{x} + \underset{\sim}{A}_i (\underset{\sim}{A}_i - \underset{\sim}{A}) \underset{\sim}{x} = \underset{\sim}{0} \ .$$

By Lemma 3.1 $C(\underset{\sim}{A}_i)$'s are independent and therefore $\underset{\sim}{A}_j \underset{\sim}{A}_i \underset{\sim}{x} = \underset{\sim}{0}$, $\forall \underset{\sim}{x}$, $j \neq i$.
Hence

$$\underset{\sim}{A}_j \underset{\sim}{A}_i = \underset{\sim}{0} \ , \quad j \neq i \ .$$

Now

$$\underset{\sim}{0} = \underset{\sim}{A}^r - \underset{\sim}{A} = \sum_{i=1}^k (\underset{\sim}{A}_i^r - \underset{\sim}{A}_i) \ .$$

Noting $(\underset{\sim}{A}_i^r - \underset{\sim}{A}_i) \underset{\sim}{x} \in C(\underset{\sim}{A}_i)$, $\forall \underset{\sim}{x}$, we get

$$\underset{\sim}{A}_i^r = \underset{\sim}{A}_i , \quad i = 1, \ldots, k \ . \quad Q.E.D.$$

We have assumed $r \geq 3$ above because (iii') is not needed in the
case $r = 2$. Anderson and Styan (1980, their Theorem 3.3) replaces
(iii') by $\underset{\sim}{A}_i \underset{\sim}{A}^{r-2} = \underset{\sim}{A}^{r-2} \underset{\sim}{A}_i$, for example. This condition is nice in the
sense that it reduces to the null condition when $r = 2$. On the other hand,
(iii') is desirable in connection with simultaneous diagonalizability
(Theorem 4.3 in section 4) and actually easier to check for $r > 3$.

Our proof of sufficiency is quite general and it yields the follow-
ing result almost without changes.

<u>Theorem 3.3.</u> Let $P(x)$ be any polynomial of deg $P \geq 2$. Suppose

        (i) $P(\underset{\sim}{A}) = \underset{\sim}{0}$,

        (ii) rank $\underset{\sim}{A} = \sum_{i=1}^k$ rank $\underset{\sim}{A}_i$,

        (iii) $\underset{\sim}{A} \underset{\sim}{A}_i = \underset{\sim}{A}_i \underset{\sim}{A}$.

Then

$$A_{\sim i}A_{\sim j} = \underset{\sim}{0} \quad \text{and} \quad \begin{cases} P(A_{\sim i}) = \underset{\sim}{0}, & i = 1,\ldots,k, \quad \text{if} \quad P(0) = 0, \\ \\ A_{\sim i}\,P(A_{\sim i}) = \underset{\sim}{0}, & i = 1,\ldots,k, \quad \text{if} \quad P(0) \neq 0. \end{cases}$$

Proof. As in the proof of Theorem 3.2 we get $A_{\sim i}A_{\sim j} = \underset{\sim}{0}$, $i \neq j$.

Therefore $\underset{\sim}{0} = P(A_{\sim}) = \sum_{i=1}^{k} P(A_{\sim i})$ if $P(0) = 0$. Then

$$\underset{\sim}{0} = \sum_{i=1}^{k} A_{\sim i}\,Q(A_{\sim i})$$

where $Q(x) = P(x)/x$. By independence of $C(A_{\sim i})$'s we get

$A_{\sim i}\,Q(A_{\sim i}) = \underset{\sim}{0}$, $i = 1,\ldots,k$. Therefore $P(A_{\sim i}) = A_{\sim i}\,Q(A_{\sim i}) = \underset{\sim}{0}$. If

$P(0) \neq 0$, let $R(x) = xP(x)$. Then $R(0) = 0$, $R(A_{\sim}) = \underset{\sim}{0}$, and by

the above argument $R(A_{\sim i}) = A_{\sim i}\,P(A_{\sim i}) = \underset{\sim}{0}$, $i = 1,\ldots,k$. Q.E.D.

When $P(0) = 0$ the converse of this is true except for (ii).
A partial converse will be given in Corollary 4.2.

4. Decomposition to Projections

The decomposition discussed in the previous section may be called the

Cochran-type decomposition. Here we want to investigate another type

of decomposition which we call the decomposition to projections.[3] For a

diagonalizable[4] matrix such decomposition was discussed at the end of

---

[3] It is often called "spectral decomposition, as in the case of orthogonal projections.

[4] A more "mathematical" term is "semisimple".

section 1. First we derive a simple criterion which guarantees the diagonalizability of a matrix. After that the relation of two types of the above decompositions will be discussed and the main theorem of this section (Theorem 4.4) asserts that two types of projections commute. Theorem 4.4 will be applied to generalize Cochran's theorem to sum of weighted χ^2 variables.

To discuss the diagonalizability of matrices we rely on the theory of Jordan canonical form and minimal polynomial. For a $p \times p$ matrix $\underset{\sim}{A}$ let $\lambda(\underset{\sim}{A}) = \{\lambda_1, \ldots, \lambda_k\}$ and let m_i be the multiplicity of eigenvalue λ_i, $i = 1, \ldots, k$ $(\sum_{i=1}^{k} m_i = p)$. $E(\lambda_i) = N(\underset{\sim}{A} - \lambda_i \underset{\sim}{I})$ is called the eigenspace associated with λ_i. $E^W(\lambda_i) = N([\underset{\sim}{A} - \lambda_i \underset{\sim}{I}]^{m_i})$ is called the wide-sense eigenspace associated with λ_i. A polynomial $\phi_{\underset{\sim}{A}}(\cdot)$ is called the minimal polynomial of $\underset{\sim}{A}$ if $\phi_{\underset{\sim}{A}}$ is the lowest degree polynomial such that $\phi_{\underset{\sim}{A}}(\underset{\sim}{A}) = \underset{\sim}{0}$. The theory of Jordan canonical forms tells us the following: $E^W(\lambda_i)$, $i = 1, \ldots, k$, are independent, $\dim E^W(\lambda_i) = m_i$, and $\sum_{i=1}^{k} E^W(\lambda_i) = X$ (the whole space); there exists a nonsingular $\underset{\sim}{T} = (\underset{\sim}{T}_1, \ldots, \underset{\sim}{T}_k)$ with $\underset{\sim}{T}_i \colon p \times m_i$ consisting of a basis of $E^W(\lambda_i)$ such that

$$
\underset{\sim}{T}^{-1} \underset{\sim}{A} \underset{\sim}{T} = \begin{pmatrix} \underset{\sim}{C}_1 & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{C}_2 & \cdots & \underset{\sim}{0} \\ \vdots & \vdots & & \vdots \\ \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{C}_k \end{pmatrix}
$$

where $\underset{\sim}{C}_i$ is an $m_i \times m_i$ upper triangular matrix; furthermore $\underset{\sim}{C}_i$ is of the form

12

$$C_i = \begin{pmatrix} B_{i1} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & B_{i\ell_i} \end{pmatrix}$$

where B_{ij} is a $\beta_{ij} \times \beta_{ij}$ matrix which has the form

$$\begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix} \quad .$$

The minimal polynomial ϕ_A turns out to be

(4.1) $$\phi_A(x) = \Pi_{i=1}^{k}(x-\lambda_i)^{\beta_i}$$

where $\beta_i = \max_j \{\beta_{ij}\}$. See Shilov (1971, Chapter 6) for example.

If $\beta_i = 1$ then $C_i = \lambda_i I_{m_i}$. Therefore we have

Fact 3.1.

The following conditions are equivalent:

(i) A is diagonalizable.

(ii) $\beta_i = 1$, $i = 1,\ldots,k$.

(iii) $E(\lambda_i) = E^W(\lambda_i)$, $i = 1,\ldots,k$.

(iv) $\sum_{i=1}^{k} E(\lambda_i) = X$.

These conditions are rather hard to check for particular matrices. Therefore it is desirable to have a sufficient condition for the diagonalizability. Here is a strong but fairly general sufficient condition.

Lemma 4.1. Let $P(x)$ $(P\not\equiv 0)$ be a polynomial such that $P(x) = 0$ has no multiple root. Suppose $P(\underset{\sim}{A}) = \underset{\sim}{0}$; then $\underset{\sim}{A}$ is diagonalizable.

Proof. Let

$$P(x) = Q(x) \; \phi_{\underset{\sim}{A}}(x) + R(x)$$

where $\deg R < \deg \phi_{A}$. Then $P(\underset{\sim}{A}) = Q(\underset{\sim}{A}) \; \phi_{\underset{\sim}{A}}(\underset{\sim}{A}) + R(\underset{\sim}{A})$ and we have $R(\underset{\sim}{A}) = \underset{\sim}{0}$. But minimality of ϕ_{A} implies $R(x) \equiv 0$. Therefore $P(x) = Q(x) \; \Pi_{i=1}^{k}(x-\lambda_{i})^{\beta_{i}}$ by (4.1). By assumption about P we get $\beta_{i} = 1$, $i = 1,\ldots,k$, and P is diagonalizable. Q.E.D.

Applying this criterion we obtain the following theorem.

Theorem 4.1. (Decomposition to Projections) Let $P(x)$ $(P\not\equiv 0)$ be a polynomial such that $P(x) = 0$ has no multiple root. Suppose $P(\underset{\sim}{A}) = \underset{\sim}{0}$; then $\underset{\sim}{A}$ is diagonalizable and

$$(4.2) \qquad \underset{\sim}{A} = \lambda_{1} \underset{\sim}{H}_{1} + \ldots + \lambda_{k} \underset{\sim}{H}_{k}$$

where $\lambda(\underset{\sim}{A}) = \{\lambda_{1},\ldots,\lambda_{k}\}$, $\underset{\sim}{H}_{i}$'s are independent projections and $\sum_{i=1}^{k} \underset{\sim}{H}_{i} = \underset{\sim}{I}$. This expression is unique.

Proof. By Lemma 4.1 $\underset{\sim}{A}$ is diagonalizable. Decomposition (4.2) is given in (2.3). Q.E.D.

Corollary 4.1. Let $\underset{\sim}{A}$ be r-potent, i.e., $\underset{\sim}{A}^{r} = \underset{\sim}{A}$. Then

$$(4.3) \quad \underset{\sim}{A} = \underset{\sim}{H}_1 + \omega\underset{\sim}{H}_2 + \ldots + \omega^{r-2}\underset{\sim}{H}_{r-1}$$

$$= \underset{\sim}{H}_1 + \omega\underset{\sim}{H}_2 + \ldots + \omega^{r-2}\underset{\sim}{H}_{r-1} + 0 \cdot \underset{\sim}{H}_r$$

where $\omega = e^{i2\pi/(r-1)}$ and $\underset{\sim}{H}_i$'s are independent projections. This expression is unique.

Proof. Let $P(x) = x^r - x = x(x-\omega) \ldots (x-\omega^{r-2})(x-1)$. Then $P(x)$ has no multiple root. Noting that $\lambda(\underset{\sim}{A}) \subset \{1,\omega,\ldots,\omega^{r-2},0\}$ (4.3) follows from (4.2). Q.E.D.

This decomposition has the following intuitive meaning. First note $C(\underset{\sim}{A}) = \sum_{i=1}^{r-1} C(\omega^{i-1}\underset{\sim}{H}_i) = \sum_{i=1}^{r-1} C(\underset{\sim}{H}_i)$. Now $\underset{\sim}{A}^r = \underset{\sim}{A}$ implies $C(\underset{\sim}{A}) = C(\underset{\sim}{A}^r)$ $= C(\underset{\sim}{A}^{2r-1}) = \ldots$. On the other hand, $C(\underset{\sim}{A}) \supset C(\underset{\sim}{A}^2) \supset \ldots$, therefore, $C(\underset{\sim}{A}) = C(\underset{\sim}{A}^k)$, $k = 2,3,\ldots$. Let $\underset{\sim}{x} \in C(\underset{\sim}{A})$ then $\underset{\sim}{x} = \underset{\sim}{A}\underset{\sim}{y}$. Then $\underset{\sim}{A}^{r-1}\underset{\sim}{x} = \underset{\sim}{A}^r\underset{\sim}{y} = \underset{\sim}{A}\underset{\sim}{y} = \underset{\sim}{x}$. Therefore, $\underset{\sim}{A}$ can be interpreted as "rotating" $C(\underset{\sim}{A})$ completely in $r - 1$ steps. Actually (4.3) tells us that $\underset{\sim}{A}$ divides $C(\underset{\sim}{A})$ into $r - 1$ independent spaces $C(\underset{\sim}{H}_i)$, $i = 1,\ldots,r - 1$, and rotates $C(\underset{\sim}{H}_i)$ by the angle $(i-1)2\pi/(r-1)$.

At this point we have two decompositions: Cochran-type decomposition (Theorem 3.3) and decomposition to projections (Theorem 4.1). It is conceivable that these two decompositions are "independent" and the order of applying them is interchangeable.

Theorem 4.2. (Commutativity)

Let $P(x)$ ($P\not\equiv 0$) be a polynomial with no multiple root. Let $\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{A}_i$. Suppose

$$\text{(i)} \quad P(\underset{\sim}{A}) = \underset{\sim}{0} \;,$$

$$\text{(ii)} \quad \text{rank } \underset{\sim}{A} = \sum_{i=1}^{k} \text{rank } \underset{\sim}{A}_i \;,$$

$$\text{(iii)} \quad \underset{\sim\sim}{A}\underset{\sim}{A}_i = \underset{\sim}{A}_i\underset{\sim}{A} \quad i = 1,\ldots,k \;.$$

Then there exist independent projections $\underset{\sim}{H}_{11},\ldots,\underset{\sim}{H}_{k1},\ \underset{\sim}{H}_{12},\ldots\underset{\sim}{H}_{k2},\ldots,$ $\underset{\sim}{H}_{1\ell},\ldots,\underset{\sim}{H}_{k\ell}$ such that

$$\underset{\sim}{A}_1 = \lambda_1 \underset{\sim}{H}_{11} + \lambda_2 \underset{\sim}{H}_{12} + \ldots + \lambda_\ell \underset{\sim}{H}_{1\ell} \;,$$

$$\underset{\sim}{A}_2 = \lambda_1 \underset{\sim}{H}_{21} + \lambda_2 \underset{\sim}{H}_{22} + \ldots + \lambda_\ell \underset{\sim}{H}_{2\ell} \;,$$

$$\vdots$$

$$\underset{\sim}{A}_k = \lambda_1 \underset{\sim}{H}_{k1} + \lambda_2 \underset{\sim}{H}_{k2} + \ldots + \lambda_\ell \underset{\sim}{H}_{k\ell} \;,$$

$$\underset{\sim}{A} = \lambda_1 \underset{\sim}{H}_1 + \lambda_2 \underset{\sim}{H}_2 + \ldots + \lambda_\ell \underset{\sim}{H}_\ell \;,$$

(4.4)

where $\underset{\sim}{H}_j = \sum_{i=1}^{k} \underset{\sim}{H}_{ij}$ and $\{\lambda_1,\ldots,\lambda_\ell\} = \lambda(\underset{\sim}{A}) - \{0\}$. These expressions are unique.

We first prove the following essential

Theorem 4.3. (Simultaneous Diagonalization)

Let $\underset{\sim}{A},\underset{\sim}{B}$ be diagonalizable. Then

$$\underset{\sim\sim}{AB} = \underset{\sim\sim}{BA}$$

if and only if $\underset{\sim}{A},\underset{\sim}{B}$ are simultaneously diagonalizable.

<u>Proof.</u>

Sufficiency: Let $\underset{\sim}{A} = \underset{\sim}{T}\underset{\sim}{D_1}\underset{\sim}{T}^{-1}$, $\underset{\sim}{B} = \underset{\sim}{T}\underset{\sim}{D_2}\underset{\sim}{T}^{-1}$, then $\underset{\sim}{A}\underset{\sim}{B} = \underset{\sim}{T}\underset{\sim}{D_1}\underset{\sim}{D_2}\underset{\sim}{T}^{-1}$ $= \underset{\sim}{T}\underset{\sim}{D_2}\underset{\sim}{D_1}\underset{\sim}{T}^{-1} = \underset{\sim}{B}\underset{\sim}{A}$.

Necessity: Let $\lambda(\underset{\sim}{A}) = \{\lambda_1,\ldots,\lambda_k\}$, $\lambda(\underset{\sim}{B}) = \{\mu_1,\ldots,\mu_\ell\}$ and $\underset{\sim}{A} = \sum_{i=1}^{k} \lambda_i \underset{\sim}{H}_i^A$, $\underset{\sim}{B} = \sum_{j=1}^{\ell} \mu_j \underset{\sim}{H}_j^B$. Suppose $\underset{\sim}{x} \in E(\lambda_i) = C(\underset{\sim}{H}_i^A)$. Then $\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{x} = \underset{\sim}{B}\underset{\sim}{A}\underset{\sim}{x}$ $= \lambda_i\underset{\sim}{B}\underset{\sim}{x}$. Therefore $\underset{\sim}{B}\underset{\sim}{x} \in E(\lambda_i)$ or $C(\underset{\sim}{B}\underset{\sim}{H}_i^A) \subset C(\underset{\sim}{H}_i^A)$. Noting $\underset{\sim}{I} = \sum_{i=1}^{k} \underset{\sim}{H}_i^A$ we get

$$0 = \underset{\sim}{B} - \underset{\sim}{B} = \sum_{i=1}^{k} (\underset{\sim}{H}_i^A\underset{\sim}{B} - \underset{\sim}{B}\underset{\sim}{H}_i^A) .$$

Therefore $\underset{\sim}{H}_i^A\underset{\sim}{B} = \underset{\sim}{B}\underset{\sim}{H}_i^A$. By Corollary 4.1 $\underset{\sim}{H}_i^A$ is diagonalizable , and repeating the same argument we obtain

$$\underset{\sim}{H}_i^A \underset{\sim}{H}_j^B = \underset{\sim}{H}_j^B \underset{\sim}{H}_i^A \text{ for any } i \text{ and } j .$$

Let $\underset{\sim}{H}_{ij} = \underset{\sim}{H}_i^A \underset{\sim}{H}_j^B$. Then

$$\underset{\sim}{H}_{ij}^2 = \underset{\sim}{H}_i^A \underset{\sim}{H}_j^B \underset{\sim}{H}_i^A \underset{\sim}{H}_j^B = (\underset{\sim}{H}_i^A)^2 (\underset{\sim}{H}_j^B)^2 = \underset{\sim}{H}_i^A \underset{\sim}{H}_j^B = \underset{\sim}{H}_{ij} ;$$

similarly

$$\underset{\sim}{H}_{ij} \underset{\sim}{H}_{i'j'} = 0 \text{ for } i \neq i' \text{ or } j \neq j' .$$

It follows that $\underset{\sim}{H}_{ij}$, $i = 1,\ldots,k$, $j = 1,\ldots,\ell$, are independent projections such that $\underset{\sim}{I} = \sum_{i,j} \underset{\sim}{H}_{ij}$. By Cochran's theorem $\sum_{i,j} \text{rank } \underset{\sim}{H}_{ij} = p$. Now taking appropriate bases from $C(\underset{\sim}{H}_{ij})$'s and letting $\underset{\sim}{T}$ as in the proof of Proposition 2.3, it is easy to see that $\underset{\sim}{A},\underset{\sim}{B}$ are simultaneously diagonalized by $\underset{\sim}{T}$ and $\underset{\sim}{A},\underset{\sim}{B}$ can be written as

$$\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{\lambda}_i \sum_{j=1}^{\ell} \underset{\sim}{H}_{ij} \ ,$$

(4.5)

$$\underset{\sim}{B} = \sum_{j=1}^{\ell} \mu_j \sum_{i=1}^{k} \underset{\sim}{H}_{ij} \ .$$

Now $C(\underset{\sim}{H}_{ij}) \subset C(\underset{\sim}{H}_i^A) \cap C(\underset{\sim}{H}_j^B) = E(\lambda_i) \cap E(\mu_j)$. On the other hand if $\underset{\sim}{x} \in E(\lambda_i) \cap E(\mu_j)$ then $\underset{\sim}{H}_{ij}\underset{\sim}{x} = \underset{\sim}{H}_i^A \underset{\sim}{H}_j^B \underset{\sim}{x} = \underset{\sim}{H}_i^A \underset{\sim}{x} = \underset{\sim}{x} \in C(\underset{\sim}{H}_{ij})$. Therefore $C(\underset{\sim}{H}_{ij}) = E(\lambda_i) \cap E(\mu_j)$. Q.E.D.

<u>Remark</u>: It is immediate to generalize the above proof to more than two diagonalizable matrices commuting with each other. For more standard proofs see Satake (1975, Sec. 4.2, ex. 1), Greub (1975, Sec. 13.27).

Now we are in a position to prove Theorem 4.2.

<u>Proof of Theorem 4.2</u>. The case $\deg P = 1$ reduces to Cochran's Theorem. Suppose $\deg P \geq 2$, then by Theorem 3.3 $\underset{\sim}{A}_i\underset{\sim}{A}_j = \underset{\sim}{0}$, $i \neq j$, and $P(\underset{\sim}{A}_i) = \underset{\sim}{0}$ or $\underset{\sim}{A}_i P(\underset{\sim}{A}_i) = \underset{\sim}{0}$. By Lemma 4.1 $\underset{\sim}{A}_i$'s can be diagonalized. Let $\{\mu_{i1},\ldots,\mu_{ih_i}\} = \lambda(\underset{\sim}{A}_i) - \{0\}$. We want to show

$$\lambda(\underset{\sim}{A}_i) - \{0\} \subset \lambda(\underset{\sim}{A}) - \{0\}$$

and

$$E(\mu_{ij}) \subset E(\lambda_\alpha)$$

where

$$\lambda_\alpha = \mu_{ij} \ .$$

Suppose $\underset{\sim}{x} \in E(\mu_{ij})$. Then

$$\mu_{ij} \underset{\sim}{A} \underset{\sim}{x} = \underset{\sim}{A} \underset{\sim}{A}_i \underset{\sim}{x} = \underset{\sim}{A}_i^2 \underset{\sim}{x} = (\mu_{ij})^2 \underset{\sim}{x} \ .$$

Dividing by μ_{ij} we get $\underset{\sim}{A}\underset{\sim}{x} = \mu_{ij} \underset{\sim}{x}$; hence $\mu_{ij} \in \lambda(\underset{\sim}{A}) - \{0\}$ and if $\mu_{ij} = \lambda_\alpha$ then $E(\mu_{ij}) \subset E(\mu_\alpha)$.

Summarizing this we obtain now from Theorem 4.2

$$\underset{\sim}{A} = \lambda_1 \underset{\sim}{H}_1 + \ldots + \lambda_\ell \underset{\sim}{H}_\ell \, ,$$

$$\underset{\sim}{A}_1 = \lambda_1 \underset{\sim}{H}_{11} + \ldots + \lambda_\ell \underset{\sim}{H}_{1\ell} \, ,$$

$$\vdots$$

$$\underset{\sim}{A}_k = \lambda_1 \underset{\sim}{H}_{k1} + \ldots + \lambda_\ell \underset{\sim}{H}_{k\ell} \, ,$$

where

(i) $\underset{\sim}{H}_1, \ldots, \underset{\sim}{H}_\ell$ are independent projections,

(ii) for fixed i $\underset{\sim}{H}_{i1}, \ldots, \underset{\sim}{H}_{i\ell}$ are independent projections, and

(iii) $E(\mu_{i\alpha}) = C(\underset{\sim}{H}_{ij}) \subset C(\underset{\sim}{H}_j) = E(\lambda_j)$, $i = 1, \ldots, k$, where

$\mu_{i\alpha} = \lambda_j$.

Now

$$\underset{\sim}{A} \, \underset{\sim}{x} = \sum_{j=1}^{\ell} \lambda_j \underset{\sim}{H}_j \, \underset{\sim}{x} \, ,$$

$$\sum_{i=1}^{k} \underset{\sim}{A}_i \, \underset{\sim}{x} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \lambda_j \underset{\sim}{H}_{ij} \, \underset{\sim}{x} \, .$$

Therefore

$$(\underset{\sim}{H}_j - \sum_{i=1}^{k} \underset{\sim}{H}_{ij}) \lambda_j \, \underset{\sim}{x} = \underset{\sim}{0} \, , \ \forall \underset{\sim}{x} \, ,$$

and we get

(4.6) $$\underset{\sim}{H}_j = \sum_{i=1}^{k} \underset{\sim}{H}_{ij} \, .$$

Let $\underset{\sim}{H} = \sum_{j=1}^{\ell} \underset{\sim}{H}_j = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \underset{\sim}{H}_{ij}$. Now

$$\text{rank } \underset{\sim}{A} = \sum_{j=1}^{\ell} \text{rank } \underset{\sim}{H}_j = \text{rank } \underset{\sim}{H} \, ,$$

$$\text{rank } \underset{\sim}{A}_i = \sum_{j=1}^{\ell} \text{rank } \underset{\sim}{H}_{ij} \, .$$

Hence

$$\sum \text{rank } \underset{\sim}{A}_i = \text{rank } \underset{\sim}{A}$$

implies

$$\sum\sum_{i,j} \text{rank } \underset{\sim}{H}_{ij} = \text{rank } \underset{\sim}{H} .$$

Therefore by Cochran's Theorem (Theorem 3.1) $\underset{\sim}{H}_{ij}$'s are independent projections. From the construction the expressions are unique. Q.E.D.

Corollary 4.2. (Converse to Theorem 3.3) Suppose $P(x)$ has no multiple root and $P(0) = 0$. Let $\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{A}_i$. If

(i) $\underset{\sim}{A}_i \underset{\sim}{A}_j = \underset{\sim}{0}$, $i \neq j$,

(ii) $P(\underset{\sim}{A}_i) = \underset{\sim}{0}$, $i = 1,\ldots,k$,

then

(iii) $P(\underset{\sim}{A}) = \underset{\sim}{0}$,

(iv) $\underset{\sim}{A} \underset{\sim}{A}_i = \underset{\sim}{A}_i \underset{\sim}{A}$,

(v) $\text{rank } \underset{\sim}{A} = \sum_{i=1}^{k} \text{rank } \underset{\sim}{A}_i$.

Proof. (iii), (iv) follow immediately from (i) and (ii). Therefore we assume (i) - (iv) instead. Then the above proof requires no change up to (4.6). Now we want to show that

$$\underset{\sim}{H}_{is} \underset{\sim}{H}_{jt} = \underset{\sim}{0}$$

whenever $i \neq j$ or $s \neq t$ without the rank condition. Fix s,t such that $s \neq t$. Let α,β be such that $\mu_{i\alpha} = \lambda_s$, $\mu_{j\beta} = \lambda_t$. If $\lambda_s \notin \lambda(\underset{\sim}{A}_i) - \{0\}$ then let $E(\mu_{i\alpha}) = \emptyset$ instead. In any case $E(\mu_{i\alpha}) \cap E(\mu_{j\beta}) \subseteq E(\lambda_s)$

20

$\cap \; E(\lambda_t) = \emptyset$. Theorem 4.3 applied to $A_{\sim i}, A_{\sim j}$ (note $A_{\sim i} A_{\sim j} = A_{\sim j} A_{\sim i} = 0$)

implies that in (4.5) the second sums are just single terms and now $s \neq t$

implies $H_{\sim is} H_{\sim jt} = 0$. Then for $i \neq j$, $0 = A_{\sim i} A_{\sim j} = \sum_{s=1}^{\ell} \lambda_s^2 H_{\sim is} H_{\sim js}$.

Noting $C(H_{\sim is} H_{\sim js}) \subset C(H_{\sim s})$ we get $H_{\sim is} H_{\sim js} = 0$. Therefore we have

shown $H_{\sim ij}$'s are independent and by Cochran's Theorem used in the other

direction we get $\text{rank } H_{\sim} = \sum_{ij} \text{rank } H_{\sim ij}$ which yields $\text{rank } A_{\sim} = \sum_{i=1}^{k}$

$\text{rank } A_{\sim i}$. Q.E.D.

Khatri (1977, Sec. 3) discusses similar results without assuming the

commutativity (iii) which is essential for our development.

Actually we can put Theorem 4.2 into a more natural and general form.

<u>Theorem 4.4</u>. (Commutativity) Let $A_{\sim} = \sum_{i=1}^{k} A_{\sim i}$. Suppose

        (i) $A_{\sim}$ is diagonalizable,

      (ii) $\text{rank } A_{\sim} = \sum_{i=1}^{k} \text{rank } A_{\sim i}$,

    (iii) $A_{\sim} A_{\sim i} = A_{\sim i} A_{\sim}$ $i = 1,\ldots,k$,

Then $A_{\sim i}$'s are simultaneously diagonalizable and (4.4) holds.

<u>Proof</u>. By Fact 4.1, $A_{\sim}$ is diagonalizable $\Leftrightarrow \phi_{A_{\sim}}$ has no multiple root.

Therefore the result follows if we let $P = \phi_{A_{\sim}}$ in Theorem 4.2 . Q.E.D.

Now looking at (4.4) we see that row-wise it gives a decomposition

to projections and column-wise it gives Cochran-type decompositions.

In addition, marginal sums add up correctly. To get $H_{\sim ij}$ we could either

(1) decompose $A_{\sim i}$ into projections, or (2) decompose $A_{\sim}$ into projections

$\underset{\sim}{H}_1, \dots, \underset{\sim}{H}_\ell$ and let $\underset{\sim}{H}_{ij} = \underset{\sim}{A}_i \underset{\sim}{H}_j$. In this sense the two decompositions commute.

Finally we give an application of Theorem 4.4 to weighted χ^2-variables.
If $\underset{\sim}{A}$ and $\underset{\sim}{A}_i$'s are real symmetric the theory developed in this section
can be interpreted entirely in real numbers and real matrices. Further-
more, diagonalization can be restricted to the ones by orthogonal matrices,
which is always possible. In statistical terms, then, Theorem 4.4 reads:

Theorem 4.5. Let $x_1, \dots, x_p$ be independent standard normal variables.
Let $Q = \underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x}$, $Q_1 = \underset{\sim}{x}'\underset{\sim}{A}_1\underset{\sim}{x}, \dots, Q_k = \underset{\sim}{x}'\underset{\sim}{A}_k\underset{\sim}{x}$ be quadratic forms in x_i's
such that

$\quad$ (i) $\underset{\sim}{A} = \sum_{i=1}^{k} \underset{\sim}{A}_i$,

$\quad$ (ii) rank $\underset{\sim}{A} = \sum_{i=1}^{k}$ rank $\underset{\sim}{A}_i$,

$\quad$ (iii) $\underset{\sim}{A}\underset{\sim}{A}_i = \underset{\sim}{A}_i\underset{\sim}{A}$, $i = 1, \dots, k$.

Then there exist independent standard normal variables $y_1, \dots, y_p$ such
that $(y_1, \dots, y_p)' = \underset{\sim}{\Gamma}(x_1, \dots, x_p)'$ for some orthogonal $\underset{\sim}{\Gamma}$ and

$$Q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_m y_m^2 ,$$

$$Q_1 = \lambda_1 y_1^2 + \dots + \lambda_{m_1} y_{m_1}^2 ,$$

$$Q_2 = \lambda_{m_1+1} y_{m_1+1}^2 + \dots + \lambda_{m_1+m_2} y_{m_1+m_2}^2 ,$$

$$\vdots$$

$$Q_k = \lambda_{m-m_k+1} y_{m-m_k+2}^2 + \dots + \lambda_m y_m^2$$

22

where λ_i's are eigenvalues of $\underset{\sim}{A}$ and $m_i = $ rank $\underset{\sim i}{A}$, $m = $ rank $\underset{\sim}{A}$.

The converse of this theorem is easy to prove. Therefore it actually gives necessary and sufficient conditions. This theorem in its matrix form is essentially proved in Luther (1965, his Theorem 1). See also Khatri (1977) for further discussion.

REFERENCES

Anderson, T. W. and Styan, G.P.H. (1980). "Cochran's theorem, rank additivity, and tripotent matrices," Technical Report No. 43. Department of Statistics, Stanford University, Stanford, CA.

Chipman, J. S. and Rao, M. M. (1964). "Projections, generalized inverses, and quadratic forms," J. Math. Anal. Appl., 9, pp. 1-11.

Cochran, W. G. (1934). "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance," Proc. Cambridge Philos. Soc., 30, pp. 178-191.

Graybill, F. A. and Marsaglia, G. (1957). "Idempotent matrices and quadratic forms in the general linear hypothesis," Ann. Math. Statist., 28, pp. 678-686.

Greub, W. (1975). Linear Algebra (4th ed.), Springer-Verlag, New York.

Khatri, C. G. (1968). "Some results for the singular normal multivariate regression models," Sankhya Ser. A, 30, pp. 267-280.

Khatri, C. G. (1977). "Quadratic forms and extension of Cochran's theorem to normal vector variables," Multivariate Analysis, IV (P. R. Khrishnaiah, ed.), North-Holland, Amsterdam, pp. 79-94.

Luther, N. Y. (1965). "Decomposition of symmetric matrices and distribution of quadratic forms," Ann. Math. Statist., 36, pp. 683-690.

Rao, C. R. and Mitra, S. K. (1971). Generalized Inverse of Matrices, Wiley, New York.

Satake, I. (1975). Linear Algebra (Translated from Japanese). Marcel Dekker, Inc., New York.

Shilov, G. E. (1971). Linear Algebra (R. A. Silverman, tr.), Prentice-Hall, Inc., New Jersey.

Takeuchi, K. (1966). Linear Algebra (in Japanese). Baifukan, Inc., Tokyo.

Taussky, O. (1966). "Remarks on a matrix theorem arising in statistics," Monatsh. Math., 70, pp. 461-464.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations,"
Steffen L. Lauritzen, May, 1974.

18. "On the Interrelationships among Sufficiency, Total Sufficiency and
Some Related Concepts," Steffen L. Lauritzen, June, 1974.

19. "Statistical Inference for Multiply Truncated Power Series Distributions,"
T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models
in the Time and Frequency Domains," T. W. Anderson, June 1975.

21. "Asymptotic Properties of Some Estimators in Moving Average Models,"
Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting,"
Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear
Discriminant Function," Chien-Pai Han, April 1976.

24. "Panels and Time Series Analysis: Markov Chains and Autoregressive
Processes," T. W. Anderson, July 1976.

25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson,
September 1976.

26. "The Recurrence Classification of Risk and Storage Processes,"
J. Michael Harrison and Sidney I. Resnick, September 1976.

27. "The Generalized Variance of a Stationary Autoregressive Process,"
T. W. Anderson and Raul P.Mentz, October 1976.

28. "Estimation of the Parameters of Finite Location and Scale Mixtures,"
Javad Behboodian, October 1976.

29. "Identification of Parameters by the Distribution of a Maximum
Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.

30. "Discrimination Between Stationary Guassian Processes, Large Sample
Results," Will Gersch, January 1977.

31. "Principal Components in the Nonnormal Case: The Test for Sphericity,"
Christine M. Waternaux, October 1977.

32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a
Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

33. "Canonical Correlations with Respect to a Complex Structure,"
    Steen A. Andersson, July 1978.

34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and
    I. Olkin, July 1978.

35. " Maximum likelihood Estimation for Vector Autoregressive Moving
    Average Models," T. W. Anderson, July 1978.

36. "Maximum likelihood Estimation of the Covariances of the Vector Moving
    Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.

37. "Efficient Estimation of a Model with an Autoregressive Signal with
    White Noise," Y. Hosoya, March 1979.

38. "Maximum Likelihood Estimation of the Parameters of a Multivariate
    Normal Distribution,"T.W. Anderson and I. Olkin, July 1979.

39. "Maximum Likelihood Estimation of the Autoregressive Coefficients and
    Moving Average Covariances of Vector Autoregressive Moving Average Models,"
    Fereydoon Ahrabi, August 1979.

40. "Smoothness Priors and the Distributed Lag Estimator," Hirotugu Akaike,
    August, 1979.

41. "Approximating Conditional Moments of the Multivariate Normal Distribution,"
    Joseph G. Deken, December 1979.

42. "Methods and Applications of Time Series Analysis – Part I: Regression,
    Trends, Smoothing, and Differencing," T.W. Anderson and N.D. Singpurwalla,
    July 1980.

43. "Cochran's Theorem, Rank Additivity, and Tripotent Matrices." T.W. Anderson
    and George P.H. Styan, August, 1980.

44. "On Generalizations of Cochran's Theorem and Projection Matrices,"
    Akimichi Takemura, August, 1980.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 44	2. GOVT ACCESSION NO. AD-A092 595	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON GENERALIZATIONS OF COCHRAN'S THEOREM AND PROJECTION MATRICES		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) AKIMICHI TAKEMURA		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE AUGUST 1980
		13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cochran's theorem, r-potent matrix, spectral decomposition, projection, diagonalizable matrix, simultaneous diagonalization

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Generalizations of Cochran's theorem including (i) nonsymmetric matrices and (ii) r-potent matrices are proved by consistent use of projection matrices. Decomposition of diagonalizable matrices into projections to eigenspaces (or spectral decomposition) and its relation to Cochran-type decomposition are studied.

DD FORM 1473 1 JAN 73   EDITION OF 1 NOV 68 IS OBSOLETE
S/N 0102-014-6601